

Applicant(s): ARIEH DON, ET AL.
Serial No.: 09/676,709
Filed: September 29, 2000

E30-048 (00123)

IN THE SPECIFICATION

Please insert the following paragraph at page 1, line 2:

This application is a continuation of U. S. Serial No. 10/400,208 filed March 27, 2003 which is a continuation of U. S. Serial No. 09/676,709 filed September 29, 2000 (now U. S. Patent No. 6,546,457).

Please amend the paragraph beginning at page 1, line 16 as follows:

This invention is generally directed to data processing systems including host processors and data stores formed at one [[ore]] or more disk array storage devices. More specifically this invention is directed to a method and apparatus for effecting an on-line, transparent reconfiguration of data on a data store, particularly a data store with striped files or logical devices.

Please amend the paragraph beginning at page 2, line 20 as follows:

Different logical devices may be stored with different structures. For example, certain RAID storage schemes and data striping structures distribute data in a single logical device over multiple physical disk drives with the objective of

Applicant(s): ARIEH DON, ET AL.
Serial No.: 09/676,709
Filed: September 29, 2000

E30-048 (00123)

achieving redundancy, load balancing, recovery and other goals. With respect to striping, many host requests to a non-striped logical device can produce [exhibit]maximum activity around a localized set of logical device addresses that tend to be concentrated on a single physical disk drive. Other physical disk drives that store other logical devices can remain relatively inactive. This uneven loading can adversely effect the operating characteristics of the disk array storage device. Striping can provide load balancing when such conditions exists. A striped logical device divides the data into a number of successive data blocks of contiguous locations on a plurality of physical disk drives. When the logical device is characterized as having tracks defined by logical cylinders and read/write heads, the data can be considered as residing on multi-track logical cylinders. A logical cylinder subset can be considered to have the capacity of a small number of cylinders in a particular physical disk drive. Data is distributed over successive physical disk drives so it is likely that multiple physical disk drives will respond to host requests and thereby balance loading.

Please amend the paragraph beginning at page 5, line 13 as follows:

Applicant(s): ARIEH DON, ET AL.
Serial No.: 09/676,709
Filed: September 29, 2000

E30-048 (00123)

One approach to shifting such data could involve taking the entire logical device out of service to reconfigure the data through a sequence of copy operations. In many situations, however, taking a logical device out of service for a considerable time period is just not acceptable. Alternatively if sufficient physical disk drives are available, all the existing data might be copied to a buffer for reconfiguration while the host continues to work with the existing data. The reconfiguration could then occur in parallel to [continued] continue interaction between the host and the existing data until the reconfiguration was complete. Then the reconfigured data could be substituted for the existing data or could be copied to replace the existing data and then reattached to the host. However, this approach requires some mechanism to track any changes a host application makes to the existing data during the reconfiguration and could require the data to be taken off-line for a considerable period of time. Moreover such a process can require significant processor resources that could adversely effect system performance even further. What is needed is a method and apparatus for enabling the reconfiguration of data concurrently with and transparently to host operations with the data being reconfigured.

Applicant(s): ARIEH DON, ET AL.
Serial No.: 09/676,709
Filed: September 29, 2000

E30-048 (00123)

Please amend the paragraph beginning at page 7, line 3 as follows:

In accordance with one aspect of this invention, the reconfiguration of a striped logical device that is distributed across a plurality of physical disk drives and that operates in response to input-output requests from a host includes the step of making a copy of the striped logical device in the original configuration concurrently with operations between the host and the logical device according to an original configuration. After the copy is isolated, the logical device is switched to its new configuration. The host thereafter immediately interacts with the logical device in its new configuration. Data is transferred from locations in the isolated copy to corresponding locations according to the new configuration concurrently with host interaction with the logical device. An input-output request from the host for non-transferred data initiates a transfer of data from a corresponding location in the copy to the location according to the new configuration identified by the input-output configuration.

Please replace the paragraph beginning at Page 11, line 18 as follows:

For purposes of explaining this invention, it is assumed that the service processor 40 has defined a logical device that

Applicant(s): ARIEH DON, ET AL.
Serial No.: 09/676,709
Filed: September 29, 2000

E30-048 (00123)

is distributed over three physical disk drives 31, 33 and 35 in six logical cylinder subsets or stripes of two logical cylinders as shown in FIGS. 1 and 2. More specifically in this configuration a logical cylinder subset 31A stores cylinders A and B of the logical device. logical cylinder subsets 33A and 35A store logical device cylinders C and D and cylinders E and F, respectively. The striping is completed by storing logical device cylinder G and H in logical cylinder subset 31B, logical device cylinders I and J on logical cylinder subset ~~[[37B]]~~ 33B and logical device cylinders K and L on logical cylinder subset 35B. If the logical device is a meta device, the physical disk drive 31 is the head element; the physical disk drive 35 is the tail element; and the physical disk drive 33 is a meta element.

Please replace the paragraph beginning at page 12, line 14 as follows:

This invention is particularly adapted when system requirements necessitate either the modification of a stripe size or the expansion of a striped logical device. FIG. 2 depicts an expansion from an original configuration, the logical device 41, to a new configuration, the logical device 42 with an additional physical disk drive 37. The physical disk drive has logical cylinder subsets 37A and 37B to provide the increased capacity. For a striped logical device the data

Applicant(s): ARIEH DON, ET AL.
Serial No.: 09/676,709
Filed: September 29, 2000 ..

E30-048 (00123)

stored in certain cylinder subsets must be relocated to maintain the striping scheme of the logical device 41 in its original configuration. In the specific expansion of FIG. 2, data on the logical cylinder subsets 31A, 33A and ~~[[35B]]~~35A, that is the data in logical cylinders A through F, need not be moved. However, the data on logical cylinder subset 31B, i.e., data in the logical cylinders G and H, must move to the logical cylinder subset 37A. Logical cylinders I and J and logical cylinders K and L must move to the logical cylinder subsets 31B and 33B, respectively. The logical cylinder subsets 35B and 37B are available for expansion as new logical cylinders M and N and logical cylinders O and P, respectively.

Please amend the paragraph beginning at page 21, line 8 as follows:

Step 106 sets all the pending split (PS) bits for the logical device cylinders in the original configuration. In terms of FIG. 2, this step sets all the PS bits for the physical disk drives 31, 33 and 35. More specifically, step 106 sets the PS bits 91 in all the tracks in the Track ID table ~~[[83]]~~74 in FIG. 3 for the cylinders corresponding to the logical cylinder subsets 31A, 31B, 33A, 33B, 35A and 35B as they contained logical cylinders A through L. This indicates

Applicant(s): ARIEH DON, ET AL.
Serial No.: 09/676,709
Filed: September 29, 2000

E30-048 (00123)

that none of the entries in the Track ID table associated with the original configuration have been transferred to the Track ID tables associated with new configuration.

Please amend the paragraphs beginning at page 22, line 15 through page 23, line 10 as follows:

Once these preliminary operations are complete, control transfers to step 113 in FIG. 4B to wait until the BCV device indicates that all its physical disk drives are synchronized. When this occurs, the BCV device mirrors or replicates the data in its original configuration using the correspondence of physical disk drives 31, 33 and 35 to physical disk drives 46, [[44]] 47 and 48, respectively, except for any data that might be contained in the write pending slots 44.

Until this point any host accesses have been handled by the data storage facility in a normal fashion according to the logical device 41 shown in FIG. 2. When synchronization occurs, step 114 disables host access to the logical device 41. However, the interval during which host access is denied will be very short because the times required to process steps 115, 116 and 117 do not depend upon the times for performing any data transfers between tracks on the different physical disk drives. Consequently, the denial of host access will be in the millisecond time domain. Interruptions to normal host activity

Applicant(s): ARIEH DON, ET AL.
Serial No.: 09/676,709
Filed: September 29, 2000

E30-048 (00123)

in such time domains is generally acceptable in the industry and does not require any interruption of input-output processes.

Please amend the paragraph beginning at page 25, line 20 as follows:

Step 135 uses a track reassignment process described in the above-identified U. S. Patent Application Serial No. 09/303,242 to test the track in the original configuration to determine if any write pendings need to be handled for this track. First step 137 transfers the Track ID table entry for the track from the original configuration location in the Track ID table 85 to the corresponding location in the Track ID table 74 for the BCV device. Then the process handles any write pending data by updating information in the Track ID table 74 to assure that the data moves from the corresponding write pending slot to the [DCH]o(B)] address in the BCV device. This process will also clear the corresponding one of the WPk bits 65 and, if appropriate, the WP bit 64 shown in FIG. [[1]]3.

Pleas amend the paragraph beginning at page 27, line 10 as follows:

As will become apparent, either the data transfer background process or I/O handling process might reset the PS

Applicant(s): ARIEH DON, ET AL.
Serial No.: 09/676,709
Filed: September 29, 2000

E30-048 (00123)

bit before the track is selected. If this occurs there is no need to process any such track further. Consequently whenever step `[[132]]_131` determines that the PS bit is not set, control also transfers directly to step 138.

Please amend the paragraph beginning at page 28, line 22 as follows:

Step 142 determines whether the corresponding PB bit in the PB header, such as the PB header `[[90]]_89` of the track ID table 74, is set. It is set at the beginning of the operation. It would not be set, for example, if the I/O request handler of FIG. 7 had previously processed an I/O request for the corresponding track.

Please amend the paragraph beginning at page 33, line 13 as follows:

Under the second scenario, the instant split background process of FIG. 5 has handled the corresponding track, but the data transfer background process of FIGS. 6A and 6B has not acted on the track. Consequently the data in the replicated logical device `[DCH]o(B)` address has been updated with any pending write operations and the data is valid. In this case steps 136 and 137 have also set IND bit and cleared the PS bit corresponding to the `[DCH]n(M)` address. The set IND bit

Applicant(s): ARIEH DON, ET AL.
Serial No.: 09/676,709
Filed: September 29, 2000

E30-048 (00123)

indicates that the data in the [DCH]n(M) address is the original data and needs to be updated from the replicated logical device. The PB bit for the corresponding [DCH]o(B) address will remain set because the data transfer background process has not acted on the track.